Depth-Image-based Facial Analysis between Age Groups and Recognition of 3D Faces

Chiranji Lal Chowdhary[1], Richa Dhanuka[1], Arpana Kumari[1], and Chandra Mouli P.V.S.S.R.[2]
[1] School of Information Technology and Engineering, VIT University, Vellore, India
Email: chiranji.lal@vit.ac.in, richa.dhanuka@yahoo.co.in, arp.vit@gmail.com
[2] School of Computing Science and Engineering, VIT University, Vellore, India
Email: chandramouli@vit.ac.in

***Abstract*—Face recognition is still an open problem. Many 2D face recognition approaches came into light to achieve high recognition rate. But these approaches are still challenged by the changes in illuminations, expressions, pose, noise, etc. A 3D face recognition technique is proposed to overcome such challenges and to enhance robustness to expression variations. Here, we compare the person at different age groups with higher recognition rate in comparison to 2D face recognition techniques. We propose a two stage procedure of 3D face recognition based on FLD (Fisher Linear Discriminant), SURF operator and depth-image. First, FLD is used on depth-image to perform recognition and then the SURF features of 2D gray images to carry out the refined recognition. Finally, our proposed work will increase the robustness in expression variations.**

***Index Terms*—speeded up robust features, fisher linear discriminant, depth image, principal component analysis.**

I. Introduction

A computer application that identifies or verifies a digital image or a video of any human is a facial recognition system. This can be done by matching that facial image [2] with available facial database. This type of recognition system is typically used in security systems. Such systems can be compared to existing biometrics like nose, fingerprint, eyes or whole face recognition systems.

Most of the face recognition algorithms identify faces by extracting features from an image of the subject's face. The size of the image, shape of the eyes, nose, and relative position of different features of the image are analyzed by the used algorithms. The features are used to search for other images with matching features. Some other algorithms normalize the gallery of face images and then compress the face data. After that, only useful data are saved in images which are needed for face detection. A probe image is then compared with the face data.

Recognition algorithms being used include Principal Component Analysis, Independent Component Analysis, Linear Discriminate Analysis, Markov model and many more. In recent years the three-dimensional object / face recognition technique is emerging and claiming to achieve high accuracies. In 3-D face recognition technique a 3-D sensor can be used to capture the shape of a face with relevant information. The information collected by the 3-D sensor is used to identify different features on the surface of a face. These features are as the contour of the face, chin, eye and nose. Important factor in 3-D facial recognition is that this is not affected by the lighting. This technique can also identify a face with other pose angles.Biometric based face recognition systems are taken as the efficient technique in biometrics. There are some risk factors like hacking or duplication possible in fingerprint devices. One more issue when device may deny to access if the finger is dirty, wet or injured. Iris recognition devices are more interfering devices. Voice recognition devices are having some conditions, e.g. if person is suffering by throat problem, that time they may fail to give correct results. So Face recognition system will be recommended even it may not reach to be perfect but this is more advanced than other techniques.

II. Literature Survey

A. Overview

Present social world is too complex so the survival of an individual depends on the interpreting visual information about the personal factors like age, identity and emotional state of another person based on that face. Facial expressions [2], facial poses, appearance of face and its illumination are having a great impact in different adverse conditions. By using these, humans can perform face identification with good robustness without knowing more about individuals. Face recognition researchers using different techniques since 1970s, and it got a significant importance in the last two decades. The reasons behind the growth for face recognition system are the possibility of wide range of applications of it and the emergence of affordable hardware. Hardware, such as digital cameras, made the capturing of high-quality and high resolution images more similar.

Many 2D based approaches such as PCA (principal Component Analysis), FLD (Fisher Linear Discriminant), ICA (Independent Component Analysis), etc, can achieve high recognition rate under some reasonable conditions. However, they are still challenged by the changes of illumination, pose and expressions. Recent progress in computation, storage device and 3D sensors make it possible to perform recognition based on 3D face data. 3D face recognition [3, 7, 8, 11] is considered to be less influenced by the changes of illumination, pose and is more robust to expression variations to some extent. Expression variation is still a big problem not

DOI: 01.IJSIP.03.01.86

only for 2D face recognition but also for 3D face recognition. In order to enhance robustness to expression variations, some effective ways have been presented. Bronstein et al. [3] used the bending-invariant canonical form in constructing expression-invariant representations of faces. It is found that 3D face recognition methods can achieve significantly high accuracy rate compared to 2D face recognition, because 3D face recognition is a modality of facial recognition methods in which the three-dimensional geometry of the human face is used. The 3D face recognition also avoids some problems of 2D face recognition algorithms such as change in lighting, wearing spectacles, different facial expressions, head orientation and make-up. The 3D models are used in improving the accuracy of image based recognition after transforming the head into knows views.

B. Related Works

The survey on state of the art in 3D describes the latest results and also, the recent research trends, showing that the variety and sophistication of algorithmic approaches explored are expanding. It is challenging to improve the accuracy in recognition, facial expression robustness and rendering. So representation of the scene is bulky and it needs good algorithms for real-time rendering and efficient representation. Many methods are using Principal Component Analysis (PCA). RajKiran Gottumukkal et al. [9] which tested the potential of PCA in modular approach and the facial image is divided into modules and then PCA is applied over each module.

A real face is described by its 3D shape and texture so it is reasonable to use geometry and color or intensity, to improve recognition reliability. This is the idea behind 3D+2D face recognition. Ashutosh Saxena et al.[8] is based on MRF to get the monocular cues from a single still image, and respectively convert it into 3D. Belhumeur et al. [5] proposed a comparative study of eigenfaces generated by PCA and fisherfaces generated by LDA to show that fisherfaces are more robust towards expressions. Kresimir Delac et al. [6] proposed a comparative study of different pattern recognition algorithms ICA, PCA, LDA which provides an excellent difference among these algorithms.

C. Methods for acquiring 3D image Data

There are various ways for representing 3D face data. Point clouds, appearance-based, triangle mesh, layering, depth image etc are different approaches for 3D image data. Out of these, Depth Images can be computed from the real world using cameras or other scanning devices.

D. Algorithms used for feature extraction

Principal component analysis (PCA) [5, 9] is a mathematical procedure used in transforming a number of correlated variables into uncorrelated variables, i.e. principal components. PCA calculates the eigenvalue decomposition of a data covariance matrix of a data matrix after mean centering the data for each attribute.LDA (Linear Discriminant Analysis) [5, 6] is an enhancement to PCA which constructs a discriminant subspace that minimizes the scatter between images of same class and maximizes the scatter between different class images. In PCA image elements considers random variables with Gaussian distribution and also minimized the second-order statistics. But, for non-Gaussian distribution, largest variances would not correspond to PCA basis vectors. Now, Independent Component Analysis (ICA) minimizes both second-order and higher-order dependencies in the input data. ICA attempts to find the basis along which the data (when projected onto them) are statistically independent.
SIFT Operator (Scale Invariant Feature Transform) [4, 1, 10, 8] was proposed for extracting distinctive invariant features from images that can be invariant to image scale and rotation. SURF (Speeded-Up Robust Features) algorithm was first proposed by Bay in 2006 [10]. Essentially, it is the improvement of SIFT algorithm. At present, SURF has been applied to image registration, camera calibration and object recognition. Many SIFT or SURF-based researches have been carried out in face recognition. In order to enhance robustness to expression variations, we propose a two-stage procedure of 3D face recognition based on depth image and SURF Operator. We use FLD method on the depth image to perform coarse recognition, and then we extract the SURF features of the 2D gray images to carry out the refined recognition.

III. Proposed System

We propose a 3D Face recognition and facial age analysis system which includes two steps.

- First, carry out coarse recognition using FLD on the depth images to increase the robustness and recognition rate.
- FLD is an enhancement to PCA. It constructs a discriminant subspace that minimizes the scatter between images of same class and maximizes the scatter between different class images.
- Let $X_1, X_2, ..., X_c$ be the face classes in the database and let each face class $X_i, i = 1,2,...c$ has k facial images X_j, $j = 1,2,...,k.$
- The average face is calculated as

$$A = \frac{1}{kc}\sum_{i=1}^{c}\sum_{j=1}^{k} X_{ij}$$

- Each face differs from the average face by the vector

$$Y_{ij} = X_{ij} - A$$

- The Covariance matrix C is obtained as:

$$C = \frac{1}{kc}\sum_{i=1}^{c}\sum_{j=1}^{k} Y_{ij}.Y_{ij}^{T}.$$

- The Eigenvectors are calculated from this covariance matrix.
- We compute the mean image μ_i of each class X_i as:

$$\mu_i = \frac{1}{k}\sum_{j=1}^{k} x_j$$

DOI: 01.IJSIP.03.01.86

• Now, the mean image μ of all the classes in the database can be calculated as:

$$\mu = \frac{1}{c}\sum_{i=1}^{c}\mu_i$$

• We calculate within-class scatter matrix as:

$$S_w = \sum_{i=1}^{c}\sum_{X_k \in X_i}(X_k - \mu_i)(X_k - \mu_i)^T$$

• We calculate the between-class scatter matrix as:

$$S_B = \sum_{i=1}^{c} N_i(\mu - \mu_i)(\mu - \mu_i)^T$$

• We calculate the projection matrix as:

$$W_{LDA} = S_W . S_B$$

• We extract the feature vectors as:

$$I_{ij} = W_{LDA}^T W_{PCA}^T x_{ij}$$

• Secondly, extract the SURF features of the 2D gray images corresponding to 3D faces obtained by previous step, to perform the refined recognition.

Distance = sum ((Finaldatatesting – Finaldatatraining).^2).^0.5
M = mean (Distance)
L = min (Distance)
Percentage = ((M – L)/M)* 100

A. Assumptions

3D Face recognition should overcome the differences which occur in faces due to viewing angle, facial expressions, illumination etc. In this project we are concerned about the differences related to facial expressions, so viewing angle is assumed to be constant for all the images. We are taking a facial expression database consisting of training images and testing images for 3D face recognition system.

We have used bmp files for both the training and testing set of images. We have assumed that the image format remains constant throughout the process. Though we have used bmp files only, it can work well with other image formats also. The size of all the images is assumed to be constant as variation in size will lead to resizing of all the images to a constant value. The facial images which we have used are of high quality and thus noise is ignored. If noisy image is to be used, then we need to filter it.

B. Architecture Specification

This 3D Face recognition system consists of three modules namely acquisition of depth image, feature extraction using FLD method based on depth image, recognition using SURF operator.

Acquisition of depth image:

3D face data is obtained from the 2D image data by the use of layering. Layering provides a common colormap to all the images. The main concept used for layering is that in any image, the part which is closer to the viewer is represented by light colors as compared to the portion which is farther away. This concept is used for the conversion of 2D image into 3D perceived image. The brighter part of the image represents lesser depth as compared to the darker part.

Feature extraction using FLD method based on depth image:

Though the dimensions of depth image used in this work is reduced after pretreatment, the dimensions is still too high to use FLD directly so we are combining PCA with FLD to solve this problem. The feature vectors are generated after applying PCA and FLD respectively over the images. The process of PCA and FLD has been already described above. These vectors are used for the similarity measurement.

Recognition using SURF operator:

In this module we will compute the recognition rate by calculating the similarity between training images and testing image using SURF operator. The similarity is calculated by calculating the distance between testing and training images.

IV. RESULTS AND DISCUSSIONS

In *FRAV3D_1* database [12], we have used three classes representing three different persons. Each class has 14 images representing different expressions of the same person within the class.

A. Output / Results

In this proposed system, the outputs will vary as per the given inputs. We are considering three different possible outputs which are given below:

Figure1. (a) 2D testing image (b) Most Similar Image

Expression variant testing:

We have taken a 2D testing image as an input shown in Figure 1(a), which is an expression variant image of the images already existing in the database. Now, this input is processed as per the proposed system and the output will show the most similar image of the given input as shown in Figure 1(b). The output will show the class to which the particular input image belongs to, as shown in Figure 2.

Figure2. Matched Class

DOI: 01.IJSIP.03.01.86

Input testing image is not present in the database:

We have taken a 2D testing image as an input shown in Figure 3(a), whose class does not exist in the database.

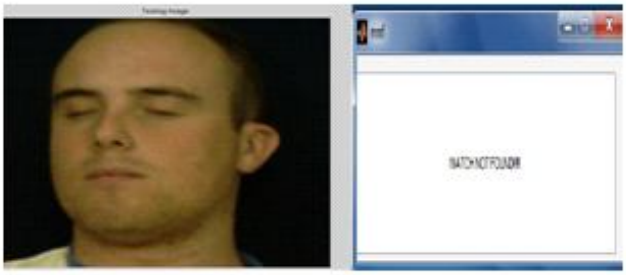

Figure3. (a) 2D Testing Image (b) Match Not Found

Now, this input is processed as per the proposed system. As this particular input image does not exist in our database, so it will pop up a window showing that the match is not available in the database as shown in Figure 3(b).

Input image for facial age analysis:

For facial age analysis, the input image will be an image of a child as shown in Figure 4(a). Now, this input is processed as per the proposed system and the output will show the similar image of elder age of the given input as shown in Figure 4(b).

Figure1. (a) Younger Image (b) Elder Image

B. Result Analysis

The algorithm was tested on *MATLAB 7.6.0*. The implementation had 14 expression samples each of 3 different classes. We performed LDA with PCA on the actual samples. There are different types of complexity which we go through while implementing the algorithm.The experiment was repeated with different number of eigenvectors.

It was observed that, as number of eigenvectors increase the recognition rate also increases as shown in Figure 5, but at the cost of computational complexity. It was observed that LDA with PCA is able to achieve higher rates of recognition than only PCA. Using only PCA implementation, the accuracy of recognition is low and inconsistent as it is unable to recognize images having different expressions and it is very much affected due to same illumination of all the images.

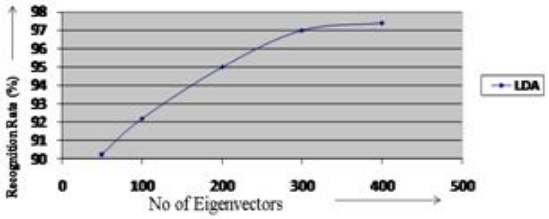

Figure1. Results of LDA for different Eigenvectors

C. Discussion

From the experiments done over the images, it is found that 3D data provides a higher rate of similarity as compared to 2D data. It can be found from Figure 6 that there is an upward shift in the graph of LDA with 3D.This shows the better similarity of the images. It can also be seen in the graph that there is less fluctuation in the similar class of LDA with 3D as compared to the similar class of 2D. But the dissimilarity among different classes also decreases in the case of LDA with 3D, which can be considered as a drawback of this experiment.

In both the cases i.e. LDA with 2D images and LDA with 3D images, the curve is smooth around the dissimilar class. So it can be concluded that 3D data does not provide much difference with 2D data in case of dissimilarity.

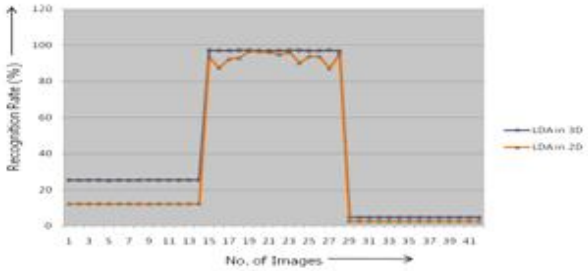

Figure6. Comparison of LDA in 3D and LDA in 2D

TABLEI. Comparison Table

Recognition Methods	Recognition Rates (%)		
	Case 1	*Case 2*	*Case 3*
PCA (2D)	100	70.22	50.06
PCA+FLD+SURF(2D)	100	91.34	27.44
PCA(3D)	100	85.57	33.86
PCA+FLD+SURF(3D)	100	97.83	0.05

Case 1: Testing image is present in database

Case 2: Testing image is an expression variant image of the training set images

Case 3: Testing image whose class is not present in the database.

Conclusions

The experiments on the database *FRAV3D_1* [12] shows that face recognition with high accuracy and recognition rate is still a challenge. Though 3D face data is more robust to expression variation, the high accuracy rate is still a challenge. The accuracy rate of 3D face recognition depends on the pretreatment of the images. There are various techniques for converting a 2D data into 3D data but all of them itself is a very broad research area. 3D face recognition technique is much more complicated as compared to that of 2D techniques.The FLD algorithm is applied over the PCA extracted feature vectors for extracting the expression

DOI: 01.IJSIP.03.01.86

invariant feature vectors. After this, the SURF operator algorithm is applied to calculate the similarity and dissimilarity. The accuracy rate is found to be 97.83%, shown in Table I.

The same procedure is applied for the age analysis also. But it is found that the recognition rate for expression variant images is high as compared to age differences. The images used for age analysis is a real time image and thus noisy, which leads to low recognition rate. Also, only PCA is applied on the images, as the database for various expressions was not used hence LDA cannot be applied, which also leads to a low recognition rate. So, a better approach can be used to get a high recognition and accuracy rate.

References

[1] Lei Yunqi, Lai Haibin, and Jiang Xutuan, "3D Face Recognition by SURF Operator based on depth image", Department of computer science, Xiamen University, IEEE, pp 240-244, 2010.

[2] Q. M. Rizvi and R. Asthana, "Facial Analysis Between Age Groups Using Distance Matrices", Tryst Technical Conference, pp 34-38, 2010.

[3] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Three-dimensional Face Recognition", Journal of Computer Vision (IJCV), Vol. 64/1, pp.5-30, August 2005.

[4] Luo Juan, Oubong Gwun, "A Comparison of SIFT, PCA-SIFT and SURF", International Journal of Image Processing (IJIP), Volume (3), Issue (4), pp 143-152, 2009.

[5] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegmanf, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", IEEE Trans. Pattern Anal. Machine Intell. 19 (7), 711-720, 1997.

[6] Kresimir Delac , Mislav Grgic, and Sonja Grgic, "A comparative study of ICA, PCA and LDA", Wiley Periodicals, Inc , Vol. 15, 252–260, 2006.

[7] Hanqi Zhuang, Teerapat Theerawong, Xin Guan, Salvatore Morgera, and Abhi Pandya, "A Method for Creating 3D Face from a 2D Face Image", Florida Conference on Recent Advances in Robotics, May 25-26, 2006.

[8] Ashutosh Saxena, Min Sun, and Andrew Y. Ng, "Make3D: Depth Perception from a Single Still Image", Association for the Advancement of Artificial Intelligence (www.aaai.org), 2008.

[9] RajKiran Gottumukkal, and Vijaya K. Asari, "An improved face recognition technique based on modular PCA approach", Pattern Recognition Letters Vol. 25 pp 429–436, 2004.

[10] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features", European Conference on Computer Vision, 1:404-417, 2006.

[11] Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel, "Expression-Invariant 3D Face Recognition", Proceedings of the 4th international conference on Audio- and video-based biometric person authentication AVBPA'03, LNCS 2688, pp. 62-70, 2003.

Chiranji Lal Chowdhary is an Assistant Professor (Sr.) and an internal part time research scholar at VIT University, Vellore. His research interests are in Image Processing and Computer Vision. He is a life member of Indian Science Congress Association (ISCA).

Richa Dhanuka is presently working in Misys Software Solutions, Bangalore. She had BTech-IT from VIT University, Vellore in 2011.

Arpana Kumari had B.Tech-IT from VIT University, Vellore in 2011. She got an offer from Sapient Corporation Private Limited, Bangalore.

Dr. Chandra Mouli P.V.S.S.R. is an Associate Professor, School of Computing Science and Engineering, VIT University, Vellore, India. He received his Ph.D. in Digital Image Processing from NIT Tiruchirappalli, India. His research interests include Digital Image Processing, Computer Vision, Pattern Recognition, Biometrics, Bio-informatics and Image Retrieval. He has published more than 13 refereed research papers in various international and national journals and conferences. He was associated with two sponsored research projects as Co-Investigator. Presently, he is working on a funded project from DRDO as Principal Investigator. He is an IEEE member and life member of Indian Society of Technical Education (ISTE). He is a reviewer for many international journals.

DOI: 01.IJSIP.03.01.86